FOR WHAT APPLICATIONS CAN PROBABILITY AND NON-PROBABILITY SAMPLING BE USED?

H. T. SCHREUDER[1], T. G. GREGOIRE[1] and J. P. WEYER[2]
[1] *USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, U.S.A.;* [2] *York University, New Haven, Connecticut, U.S.A.*

(Received 9 June 1999; accepted 28 October 1999)

Abstract. Almost any type of sample has some utility when estimating population quantities. The focus in this paper is to indicate what type or combination of types of sampling can be used in various situations ranging from a sample designed to establish cause-effect or legal challenge to one involving a simple subjective judgment. Several of these methods have little or no utility in the scientific area but even in the best of circumstances, particularly complex ones, both probabilistic and non-probabilistic procedures have to be used because of lack of knowledge and cost. We illustrate this with a marbled murrelet example.

Keywords: credibility, design-based inference, inductive logic, inferences, model-based inference, non-probability samples, probability samples, sampling protocols

1. Introduction

Environmental and ecological data are important for our understanding of the structure and functions of our ecosystems. When environmental or ecological data are obtained by sampling but the design and other aspects of the sampling protocol are unknown, it may be difficult to extrapolate these data to some larger population or cohort. Nonetheless they constitute information that has utility and value, at the very least to describe and characterize the particular sample. When the sampling strategy, that is, the combination of a sample design and an estimator of a population quantity, is known and the target population is identified, extrapolation is possible and more credible. Taking a very broad view, sampling may be probabilistic or not (a probability sample is one for which every unit in a finite population has a positive probability of selection, not necessarily equal to that of other units); it may be informative or uninformative with respect to the variable of interest; it may include response and self-selection bias and be subject to measurement error; and it may have a temporal or longitudinal structure combined with a spatial structure. Because all these factors affect estimation and inference, it is important to understand the constraints imposed by the sampling protocol on the interpretation of data. In this paper we focus upon one aspect of this multifaceted problem namely how the scope and validity of inference is impacted by sample selection. More

Environmental Monitoring and Assessment **66:** 281–291, 2001.

specifically, the titular concern is the manner in which data from probability and non-probability samples can be combined for various inferential purposes

The motivation for this study arose from historical practice and emerging need. The U.S. Forest Service (USFS) has had a long tradition of conducting probabilistic surveys e.g. in Forest Inventory and Analysis (FIA) that started in 1929 (Birdsey and Schreuder, 1992). Other surveys, particularly on National Forest (NF) lands were more often 'seat-of-the-pants' samples, or cruises, with little statistical validity. They were applied because of the prevailing view that valid surveys were impractical and could not be afforded (often true) or would yield embarrassing results. Examples are James and Schreuder (1971) where the estimates generated were embarrassing, and Schreuder *et al*. (1980) where the techniques developed were unacceptable because of costs involved. Because of the threat of lawsuits coupled with an increased awareness of the need for statistically defensible survey methods among government, industry, and environmental groups, use of probability samples is becoming mandatory (see for example Max *et al*., 1997).

Nonetheless, there remain many situations for which environmental and natural resource data are perceived to be useful even when not collected in a statistical or probabilistic manner. Moreover, much environmental data can not be collected in such a way because of expense, time constraints, or other impediments. Such data are nonetheless important and useful, providing that we can understand properly their limitations.

Currently, inventory and monitoring issues are the most litigious in the USFS. As yet, the quality of the natural resource data collected by the organization has not been challenged, but this may happen soon. As noted by Meier (1986), the Supreme Court has sanctioned formal statistical inference in a series of EEO decisions. It is only a matter of time before the role of statisticians in the legal sector is extended from cases of discrimination in employment to issues related to collecting and interpreting environmental data.

With the pervasive concern over environmental quality and forest health, survey data will be scrutinized more closely to discern dose-response and cause-effect relationships. Documenting cause-effect is very difficult as illustrated by the most noteworthy example of the last 50 yr, the association of smoking and lung cancer. In the mid 1950's, a link between smoking and lung cancer was established based on epidemiological survey data. Causality was established through verified prediction in the next 20 yr. Only recently has the causal mechanism of how smoking causes lung cancer been established (Peifer, 1997).

As we head into the information age, data on environmental and ecological systems will become increasingly more available, the need to understand the strength of these data for inference likewise becomes increasingly important.

2. Review of Literature

The process of drawing conclusions from the analysis of observed data about unobserved parameters or underlying laws is called inductive logic, one of the most controversial issues in philosophy. A discussion of the way in which data from different sources may be combined necessarily involves considering the purposes for which they are obtained. Inference, and the mechanism of inductive logic, is not limited to the comparatively narrow field of scientific and statistical inference. Nonetheless, the latter is an important sphere of activity, and a proper understanding of statistical inference is crucial to a discussion of the role of sampling in the inferential process.

Scientific inference becomes statistical inference when the connection between the unknown 'state of nature' and the observand is expressed in probabilistic terms (Dawid, 1984). There are two dominant paradigms for statistical inference from sample surveys. One approach, called model-based inference, relies on a statistical model to describe how the probability structure of the observed values depends on uncontrollable chance variables and often on other unknown nuisance variables. Such models may be based on a theoretical understanding of the procedure by which the data are generated, past experience with similar processes, or experimental techniques used. But they can be ad hoc too, chosen for ease of interpretation or analysis. The other approach, called design-based inference, relies on the probabilistic nature of the sampling. It has been the dominant paradigm since the mid-1930's as testified to by its exclusive exposition in nearly all the prominent texts on classical sampling theory and methods. We give below a capsule summary of both these modes of statistical inference deferring to Särndal (1978), Särndal *et al*. (1992), and Gregoire (1998) for a more comprehensive treatment.

Inherently, the design-based approach to statistical inference is tied to probability sampling, characteristics of which are that each element of the population has a nonzero probability of being selected and the probability of each possible sample can be deduced. The statistical behavior of estimators of a population attribute is reckoned with respect to these probabilities and the probability-weighted distribution of all possible sample estimates. It does not rely on the distribution, probabilistic or otherwise, of the attribute in the population being surveyed. One critique of the design-based approach is that samples that could have been drawn, but were not, are irrelevant: should not inference about a population parameter rightfully be based solely upon the observed sample? (Hájek, 1981). Nonetheless, the perceived objectivity of design-based inference has broad appeal and support (Hansen *et al*., 1983). The only assumption it makes is that observational units are selected at random so the validity of the inference only requires that the targeted and sampled population are the same (Koch and Gillings, 1984).

One hopes that a probability sample is representative of the survey population itself, although as Kruskal and Mosteller (1979) ably demonstrate, the word representative is subject to a wide array of interpretations. One of these is that a probabilistic sample is representative by definition. Others contend that a sample comprising the smallest n elements of a population of $N > n$ elements is hardly representative of the population, whether or not it is chosen probabilistically.

Other methods of sampling have been proposed with a view towards providing a sample that is more of a microcosm of the population. Särndal *et al*. (1992) have a short but illuminating discussion on several purposive sample selection methods, i.e. expert choice sampling, quota sampling, and balanced sampling. The latter two methods select samples from the population in ways that provide zero probabilities of selection for many units of the population. With such samples, accurate estimation of a population attribute may be possible, but an objective measure of precision is not possible within the design-based framework.

Schreuder and Alegria (1995) demonstrated the consequences for inference of using a probabilistic sample from a well defined population of interest, to estimate population totals if the probabilities of selection are unequal but unknown and, consequently, are treated as equal. This situation arises, for example, when a prior sample from a timber sale is used for a new purpose, e.g. ecological monitoring. Timber sale samples are often probabilistic but focused more heavily on strata with a greater concentration of timber than on other strata. Using the incorrect probabilities of selection introduces a design-based bias, which can be large.

It is important to understand clearly the strengths and weaknesses of purposive and random sampling. Often, purposive sampling can not be avoided, e.g., when sampling wildlife for which budgets are often minimal, the sampling process is difficult (because the animals are scattered and not stationary) and the information collected may have considerable political value because of an ‘Endangered Species Act’. Similarly, meteorological variables require very frequent measurement to be useful for purposes such as compiling growth and mortality models in forestry. At present, it is often impractical to gather such information at random sites such as FIA plots (when this is feasible, better growth and mortality models will be produced).

If a sample is selected in a non-probabilistic manner, inference may be made by specifying an underlying superpopulation model for the N-dimensional distribution of $Y = (Y_1, \ldots, Y_k, \ldots, Y_N) \subset U$, with Y_k a random variable tied to the kth element (Sarndal *et al*., 1992). If this model is called ξ, the actual finite population vector of interest $\underline{y} = (y_1, y_2, \ldots, y_k, \ldots, y_N)$ can be considered a realization of $\underline{Y}$. Assuming that we are interested in the random variable $Y = \sum_{i=1}^{N} Y_i$, then for a sample s, the realized values y_k of the random variable Y_k are observed for $k \in s$ and not observed for $k \in U - s$. Using the specification of ξ, then for an estimator $\hat{Y}$ of Y, the distribution of $\hat{Y} - Y$ can be derived for the specific sample s and the model-based mean square error of $\hat{Y} - Y$ can be obtained and estimated, leading to a model-based prediction interval for Y.

In this framework, inference may extend to parameters of the superpopulation model, and not just the particular population at hand. Hence the inference space is necessarily broader than design-based inference. Sample elements need not be chosen randomly or with known probability, so long as they are not selected on the basis of their y_k values. Because a probability structure has been assumed for the population itself, the distribution-free (strictly speaking not true since it relies on randomization distribution but not on an assumed distribution so that it is often considered distribution-free) properties of design-based inference are sacrificed. Conclusions and inferences depend on the validity of the model, which may be a serious liability if the model is not specified correctly. Conversely, with a correctly specified model, which can be viewed as a source of auxiliary information which is not fully utilized in the design-based framework, an increase in precision may be realized.

In forestry a strong linear relationship exists between tree volume (v) and diameter breast height squared times total height (d^2h). Such models, when fitted to data, often have R^2 values of 0.98 to 0.99, and the data exhibit heterogeneous errors with variance proportional approximately to $(d^2h)^2$ (Schreuder and Williams, 1997 and references therein). Yet even with such a strong relationship, there is still considerable risk in selecting units purposively and relying on the stipulated model for inference. Schreuder *et al*. (1990) showed unacceptable estimation bias for populations estimating total volume with d^2h as covariate.

The problems raised above are as important in epidemiology as in forestry and an excellent example of the issues in an epidemiological context are discussed by Greenland (1990). Overton *et al*. (1993), Cox and Piegorsch (1996), and Piegorsch and Cox (1996) have also examined ways to combine probability with nonprobability samples.

T. M. F. Smith, formerly a strong advocate of model-based inference, now seems noncommittal, viz., ‘My view is that there is no single right method of inference. All inferences are the product of man’s imagination and there can be no absolutely correct method of inductive reasoning. Different types of inference are relevant for different problems and frequently the approach recommended reflects the statistician’s background such as, science, industry, social sciences or government ... I now find the case for hard-line randomization inference based on the unconditional distribution to be acceptable ... Complete reconciliation is neither possible nor desirable. Vive la difference’ (Smith, 1994, p. 17).

Brewer (1994, 1995, 1998) discusses in detail a model-based sampling procedure called balanced sampling, first suggested by Royall (1970, 1971). This procedure is totally model-dependent and accepts only samples for which the sample moments equal the population moments for the covariates *x*, e.g., the sample mean equals the population mean. This presupposes considerable information about the population of *x*-values and, as Brewer (1994) indicates, the procedure is practicable primarily for one-time surveys. Even here, stratified sampling may be a suitable alternative. Further research is needed when balanced sampling should be used.

Schreuder and Reich (1998) discuss proposals for using models (imputation or tree models) to ‘update’ tree and plot information and suggest that while it is inappropriate for public data sets, it could be useful to specific users for updating data for their own purposes. But to quote Smith (1994, p. 34); with whom we concur: 'Democratic systems and good government require data bases of the utmost integrity'. Supplementing observational data may cause some in the public sector to question the integrity of the data and inferences stemming from their analysis.

For scientific and other users of survey data it seems imperative to understand the scope and strength of inference that can be made following surveys of various types. The motivation and objective of this paper is to further this understanding.

3. A Suite of Sampling Scenarios

In order to satisfy the objectives of the paper we consider the following hierarchy of inferences for which sample data are needed:

a. to draw inferences for the individual collecting the information only?
b. to draw inferences for an organization?
c. for inferences published in a scientific publication?
d. for inferences by the public at large, e.g. such as a public data base might be?
e. to identify possible cause-effect hypotheses?
f. to establish cause-effect relationships?

In the following we describe a series of sampling scenarios corresponding to this hierarchy and consider valid inferences that may be drawn from each.

Assume that data sets are available relating to mortality and basal area growth of trees in a particular forested region and each was acquired by different means, namely:

a. An interested party reports that during his travels through the region over the past 25 yr, he has noticed a sharp decrease in tree size and the forests seem to be less vigorous than they used to be.

Possible conclusion: In this situation the target population is unspecified and difficult to discern and the sample is subject to sample and site selection bias. The qualitative assessment of forest vigor is highly subjective. Therefore, no scientific conclusions can be made. This is a common situation in that all of us draw conclusions at times based on little or no information and, one might argue, this was a key reason why statistics impacted so profoundly on scientific inference. Basically, the only inference to be drawn from this anecdotal evidence is that there is a hypothesis to be examined that a change may have occurred (either in the person doing the observing, the climate, the forest, etc.).

b. A data set was purposively collected by representatives of forest industry or an environmental group to document a sharp increase in mortality and decline in basal area growth in natural pine stands in a region. No documentation is available of how the data were collected.

Possible conclusion: Here the population is defined to be natural pine stands. In contrast to the scenario in a), these data are derived from a defined measurement protocol rather than qualitative assessment. The purposive selection does not allow design-based inference about trends in mortality and basal area growth. Without more information about the criteria used for the purposive sampling, even model-based inference is untenable. Therefore, only very limited scientific conclusions can be drawn relating to an increase in mortality and decline in growth for the cohort comprising the stands actually observed. However, these data can serve useful descriptive purposes without any probabilistic interpretation (cf. Greenland, 1990, p. 6). One may scrutinize the data set for possible indicators of decline, which may suggest hypotheses to be examined later in a scientific setting.

c. FIA data are selected based on screening by variables such as undisturbed plots growing in natural stands (Gadbury *et al*., 1998 and references therein).

Possible conclusion: This is a probability sample of some population and any conclusions drawn are statistically valid, e.g. mortality increased and growth declined. The problem is that one has no idea of the population represented by the sample.

d. An FIA data set comprising all the plots in the region is used to demonstrate an increase in mortality and decline in basal area growth.

Possible conclusion: A statistically valid increase in mortality and decline in basal area growth can reasonably be inferred for the entire region.

e. The situation as in d) but some of the possible causes are measured too. The data may be analyzed for possible explanations for the decline (drought, beetle infestations, etc.) but no cause-effect can be established, only hypothesized.

f. Given the situation in e), but all possible causes are measured. Even this is usually not enough to establish cause-effect but at least the data may indicate what the most likely cause(s) is. If, in addition the analysis from the data show that at least 2 of the 3 criteria of Mosteller and Tukey, viz. consistency and responsiveness were met (Schreuder and Thomas, 1991; Olsen and Schreuder, 1997) than cause-effect has been established. Ideally, the third criterion: mechanism should be established too but this is often very difficult as noted earlier in the smoking-cancer issue. Consistency implies that the presence and magnitude of the effect is always associated with a minimal level of the suspected causal agent. Responsiveness is established if the symptoms are reproduced by experimentally exposing the population to the suspected causal agent. Mechanism establishes the cause-effect linkage by a step-by step approach, indicating clear understanding of what occurs (Mosteller and Tukey, 1977).

Possible conclusion: This is the ideal situation rarely met in practice. It is a long-term process and usually requires large scale surveys such as FIA, preceded or followed by careful experimentation (Smith and Sugden, 1988).

An example: We want to determine the maximum ($D_{\max}$) and preferred distance (D_{opt}). Marbled Murrelets (MM) will fly inland to establish a nest and we want to know for management purposes why these are the maximum and preferred distance. Clearly this assumes that there is a maximum distance $D_{\max}$ (very likely) and a preferred distance (D_{opt}) less likely) which may have to be estimated by assuming a model or models containing such parameters. The murrelet is a small sea bird that is listed as an endangered species. Considerable time and money is being spent on recovery efforts to increase the population. An important part of the recovery is knowing where exactly they nest and why.

MM nests are very difficult to find. A random selection approach to determine their location is not feasible. This is clearly a situation where, using volunteers and wildlife specialists in various federal and state agencies, a series of nests has to be found and their distance from the coast measured. Assume that we find n nests that are occupied or show evidence of having been occupied and their distances to the ocean.

Since this is not a probabilistic sample, we can not fit a statistical distribution such as the S_b distribution (Schreuder *et al*., 1993) to the data although we can use similar mathematical models to estimate the maximum ($D_{\max}$) and the preferred distance (D_{opt}). Clearly, our estimates of both parameters will be biased because the sampling frame does not include all the target population (the population of all MM that nest during the season sampled). For example, there may be nests at some distances from the ocean that are not detected because of thorny or impenetrable vegetation that deters observers but (perhaps even for the same reasons) may be attractive to the MM or perhaps have been vacated due to predator kill. If such nests tended to be very far from the ocean, their omission from the sample would result in an underestimate of the maximum nesting distance.

Despite the limitations of the sample described above, it can still be useful to help formulate hypotheses on cause-effect to explain the distances estimated. To illustrate this, assume that we formulate three hypotheses about factors determining the maximum distance for nesting (other equally valid hypotheses can be formulated, which is why establishing cause-effect relationships is so difficult), namely:

1. The MM can only fly this maximum distance and still supply its brood with food. This can best be established through experimentation.
2. The nest needs to be in a 'safe' place, relatively protected from predators. Evidence of this, certainly initially, can only be established through observational studies on all n observed breeding bird sites (both those occupied and those showing evidence of having been occupied) to characterize the nest environments including freedom and protection from predators. This may also require observation at apparently suitable nesting sites that are not used. Considerable data would need to be collected and analyses done before it would be possible to formulate a well-defined hypothesis such as a model testable by experimentation. Obviously, this hypothesis could be seriously affected by differences between the target and sampled populations. There may be characteristics of the non-sampled part of the target population that are essentially not observed in the sampled part.
3. A combination of 1) and 2) above. This would most likely result in one or more models including ($D_{\max}$) and (D_{opt}) to describe the behavior of the MM in locating nesting sites in terms of distance from the sea. To study this hypothesis properly is likely to require both experimentation and observational studies.

To establish whether these hypotheses are true requires considerable time and money. The political will to support this work over the long period of time required may not exist.

4. Recommendations

As discussed above, both purposive and design-based sampling have to play a role in developing an understanding of the status and function of our ecosystems. Cost and practicality dictates that. We suggest that the following recommendations be kept in mind to develop such understanding.

1. Use design based sampling when the target and actual population sampled are the same.
2. Consider purposive sampling when a decision is to be made quickly. The method is generally a lot cheaper and offers more protection against small sample sizes.
3. Use either design based or MB sampling for formulating hypotheses to be tested. Purposive sampling is usually more efficient for that purpose but can be more misleading.
4. Avoid any pretense that a purposive sample is a probabilistic, and hence representative, sample. Clearly state the assumptions made, what the sampled information can and cannot be used for, and draw inferences on that basis.

5. Use design-based sampling to identify cause-effect and then establish cause-effect through experimentation or perhaps more efficiently use experimentation to document cause-effect under controlled conditions and then attempt to infer generality through design-based sampling.

With further advances in technology, it may be possible to solve the problems more efficiently. If the MM and also its predators can be captured (at sea in the case of the MM and wherever with predators) and fitted with miniature radio transmitters, a larger sample of MM (and the predators too) or even a census from which a probabilistic sample could be drawn, could result. The sample could be probabilistic, or at least more representative for estimating the maximum and optimal distances for nesting. The information derived from the studies undertaken to establish the hypotheses could be used to plan the radio-tagging studies better and would not be wasted in waiting for the necessary technologies to become available.

Acknowledgements

We greatly appreciate reviews by Anthony R. Olsen and Geoffey B. Wood that made this a considerably improved manuscript.

References

Birdsey, R. A. and Schreuder, H. T.: 1992, An overview of forest inventory and analysis estimation procedures in the eastern U.S. – with an emphasis on components of change. USDA Forest Service RM Tech. Rep. RM-214, 11 p.

Brewer, K. R. W.: 1994, Survey Sampling Inference: some past perspectives and present prospects, *Pak. J. Stat.* **10**(1) A: 213–233.

Brewer, K. R. W.: 1995, ‘Combining Design-Based and Model-Based Inference, Ch. 30’, in: B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge and P. S. Kott (eds), *Business Survey Methods*, J. Wiley & Sons, Inc. N.Y., pp. 589–606.

Cox, L. H. and Piegorsch, W. W.: 1996, Combining environmental information. I. Environmental monitoring, measurement, and assessment. *Environmetrics* **7**, 299–308.

Dawid, A. P.: 1984, ‘Inference, Statistical: I’, in S. Kotz and N. L. Johnson (eds), *Encyclopedia of Statistical Sciences* **4**, pp. 89–105.

Fisher, R. A.: 1922, On the mathematical foundations of theoretical statistics, *Philos. Trans. Soc. Lond. A* **222**, 309–368.

Gadbury, G. L., Iyer, H. K., Schreuder, H. T. and Ueng, C. Y.: 1998, A nonparametric analysis of plot basal area growth using tree based models. USDA For. Serv. RM For. And Range Exp. Sta. RMRS-RP-2, 14 p.

Greenland, S.: 1990, Randomization, statistics, and causal Inference, *Epidemiology* **1**, 421–429.

Gregoire, T. G.: 1998, Design-based and model-based inference in survey sampling: Appreciating the difference, *Canadian J. For. Res* **28**, 1429–1447.

Hájek, J.: 1981, *Sampling From A Finite Population*, Marcel Dekker, Inc.

Hansen, M. H., Madow, W. G. and Tepping, B. J.: 1983, An evaluation of model-dependent and probability-sampling inferences in sample surveys, *Journal of American Statistical Association* **78**(384), 776–93.

James, G. A. and Schreuder, H. T.: 1971, Estimating recreation use of the San Gorgonio Wilderness, *J. For.* **68**(8), 490–493.

Koch, G. G. and Gillings, D. B.: 1984, 'Inference, Design Based vs Model Based', in S. Kotz and N. L. Johnson (eds), *Encyclopedia of Statistical Sciences* **4**, 84–88.

Kruskal, W. and Mosteller, F.: 1979, Representative sampling, III: The Current statistical literature, *Intern. Stat. Rev.* **47**, 245–265.

Max, T. A., Schreuder, H. T., Hazard, J. W., Teply, J. and Alegria, J.: 1997, The Region 6 Vegetation inventory and monitoring system, USDA For. Serv. PNW Res. Paper PNW-RP-493, 22 p.

Meier, P.: 1986, Damned liars and expert witnesses, *J. Amer. Stat. Assoc.* **81**, 269–275.

Mosteller, F. and Tukey, J. W.: 1977, *Data Analysis and Regression*, Addison-Wesley, Reading, MA.

Olsen, A. R. and Schreuder, H. T.: 1997, Perspectives on large-scale resource surveys when cause-effect is a potential issue, *Env. and Ecol. Stat.* **4**, 167–180.

Overton, J. M. C., Young, T. C. and Overton, W. S.: 1993, Using 'Found' data to augment a probabilistic sample: procedure and case study, *Env. Monitoring and Assessment* **26**, 65–83.

Peifer, M.: 1997, Cancer-beta-catenin as oncogene: The smoking gun, *Science* **275**, 1752–1753.

Piegorsch, W. W. and Cox, L. H.: 1996, Combining environmental information II: Environmental epidemiology and toxicology, *Environmetrics* **7**, 309–324.

Royall, R. M.: 1970, On finite population sampling theory under certain linear regression models, *Biometrika* **57**, 377–387.

Royall, R. M.: 1971, 'Linear Regression Models in Finite, Population Sampling Theory', in V. P. Godambe and D. A. Godambe (eds), *Foundations of Statistical Inference*, Holt, Rinehart and Winston, Toronto, pp. 259–279.

Särndal, C.-E.: 1978, Design-based and model-based inference in survey sampling, *Scandinavian J. Stat.* **5**, 27–52.

Särndal, C.-E., Swensson, B. and Wretman, J.: 1992, *Model Assisted Survey Sampling*, Springer-Verlag, NY, 694 p.

Schreuder, H. T. and Alegria, J.: 1995, Stratification and plot selection rules, misuses, and consequences. USDA Forest Service RM Forest and Range Exp. Sta. Res Note RM-RN-536, 4 p.

Schreuder, H. T. and Reich, R.: 'Data Estimation and Prediction for Natural Resources Public Data Sets', USDA For. Serv. RM Res Note (in press).

Schreuder, H. T. and Thomas, C. E.: 1991, Establishing cause-effect relationships using forest survey data, *For Sci* **37**, 1497–1525.

Schreuder, H. T. and Williams, M.S.: 1997, 'Weighted linear regression for tree volume using D2H and D2 as the dependent variables', USDA For. Serv. RM For and Range Exp Sta Res Paper RMRS-RP-6, 11 p.

Schreuder, H. T., Gregoire, T. G. and Wood, G. B.: 1993, *Sampling Methods for Multiresource Inventory*, J. Wiley & Sons, Inc., N.Y., 446 p.

Schreuder, H. T., Li, H. G. and Wood, G. B.: 1990, Model-dependent and design-dependent sampling procedures – A simulation study, USDA For. Serv RM For and Range Exp Sta Res Paper RM-291, 19 p.

Schreuder, H. T., Clerke, W. H., Barry, P., and Holland, D.: 1980, Two stage stratified sampling with regression to assess southern pine beetle damage, USDA For. Serv. SE For. Exp. Sta. Res. Paper SE-212, 6 p.

Smith, T. M. F.: 1994: 'Sample Surveys: 1975–1990; An age of Reconciliation?' *Intern. Stat. Rev.* **62**, 5–34.

Smith, T. M. F. and Sugden, R. A.: 1988, Sampling and assignment mechanisms in experiments, surveys, and observational studies, *Intern. Stat. Rev.* **56**, 165–180.